

Van artikel tot klinische praktijk:  
Praktische aanbevelingen voor kritische beoordeling van wetenschappelijke evidentie

Kristof Hoorelbeke, Elien Pieters, Laura M.S. de Putter, & Ernst H.W. Koster  
Vakgroep Experimenteel-Klinische en Gezondheidspsychologie  
Universiteit Gent

**Disclaimer:** Please note that this is the uncorrected Word version, we refer to the website of the publisher for the final PDF.

\*Correspondentieadres:

Henri-Dunantlaan 2

9000 Gent

België

Kristof.Hoorelbeke@UGent.be

- Kritische beoordeling van wetenschappelijke evidentie vormt een belangrijke vaardigheid voor evidence-based practice
- Verschillende hulpmiddelen zijn voor handen om deze vaardigheid te ontwikkelen
- We bieden concrete richtlijnen voor evaluatie van behandelonderzoek

## **Samenvatting**

In de klinische praktijk kan men niet langer voorbij aan de maatschappelijke vraag tot evidence-based werken. Dit houdt in dat hulpverleners wetenschappelijke evidentie voor een interventie dienen op te zoeken en integreren in de therapeutische praktijk. Kritische evaluatie van wetenschappelijke evidentie vormt daarbij een cruciale vaardigheid, wat inzicht vereist in onder andere wetenschappelijke onderzoekdesigns, statistische technieken en publicatiekanalen. Met dit artikel beogen we concrete richtlijnen te bieden om hulpverleners te helpen de wetenschappelijke evidentie op verschillende niveaus te evalueren en zo de kwaliteit van hun evidence-based werken te optimaliseren.

**Kernwoorden:** evidence-based practice, wetenschappelijke evidentie, kritische beoordeling, evaluatie

From publication to clinical practice:  
Practical guidelines for critical appraisal of scientific evidence

**Summary**

Clinicians are assumed to implement scientific findings into clinical practice, often referred to as evidence-based practice. This implies the need for critical appraisal skills, allowing clinicians to evaluate scientific evidence stemming from interventional studies. This requires a specific skillset, including extensive knowledge regarding research designs, statistics, and type of publications. We aim to provide specific guidelines allowing clinicians to critically appraise scientific evidence on several levels, which we believe is crucial for evidence-based practice.

**Keywords:** evidence-based practice, scientific evidence, critical appraisal, evaluation

De jaarlijkse stijging in wetenschappelijke publicaties en de toenemende uitbouw van initiatieven waarop wetenschappelijke evidentie beschikbaar gesteld wordt, stellen hulpverleners voor de uitdaging op de hoogte te blijven van recente evoluties in behandelonderzoek. Dit is nochtans van belang voor een wetenschappelijk onderbouwde praktijkvoering, waar implementatie van recente inzichten uit behandelonderzoek kan bijdragen tot betere behandeluitkomsten. Daarbij wordt in sterke mate beroep gedaan op de vaardigheid tot kritische evaluatie van wetenschappelijke bevindingen. Dit houdt ondermeer een kritische beoordeling van de kwaliteit, validiteit en toepasbaarheid van de bron in, wat heel wat expertise van hulpverleners vergt. Verschillende hulpmiddelen en indicatoren kunnen daarbij gebruikt worden om tot een relatief snelle inschatting van de kwaliteit van de evidentie te komen. Om dit proces te faciliteren, bieden wij in dit artikel specifieke richtlijnen aan voor *kritische evaluatie van wetenschappelijke evidentie*. Dergelijke evaluaties kunnen plaatsvinden op verschillende niveaus, waaronder: (a) *type bron*; (b) *wetenschappelijk tijdschrift*; (c) *impact van de publicatie binnen de onderzoekslijn*; en (d) *corpus van het manuscript*. Hieronder voorzien we aandachtspunten voor de evaluatie van publicaties op elk van deze niveaus.

### **Evaluatie op basis van het type bron**

Afhankelijk van de vraag zal gebruik gemaakt worden van gefilterde of ongefilterde bronnen (voor een overzicht van de relatie tussen type bron en type onderzoeksvraag, zie APA, 2006). Los van de inhoudelijke kwaliteit, kunnen deze bronnen hiërarchisch geordend worden op basis van kracht van evidentie, waarbij voor de vraag naar behandel-effectiviteit een belangrijke rol weggelegd is voor klinische trials die gebruikmaken van randomisering en adequate controlecondities (RCT's). Tabel 1 biedt hiertoe een samenvatting van de ordening van de Australian National Health and Medical Research Council (2015), wat in grote mate overeenkomt met de ordening van de Oxford Centre for Evidence-Based Medicine Working Group (2011) voor de vraag rond behandel-effectiviteit. Daarbij dient echter steeds rekening gehouden te worden met de kwaliteit van de individuele studies en de gerapporteerde resultaten op basis waarvan de bronnen al dan niet sterke evidentie kunnen bieden voor een bepaalde interventie (Guyatt et al., 2008; Tolin, McKay, Forman, Klonsky & Thombs, 2015). Men spreekt over '*zeer sterke evidentie*' indien er: (a) in hoogkwalitatieve studies een klinisch

betekenisvol effect van de interventie gerapporteerd wordt op de aan te pakken symptomatologie alsook op (b) functionele uitkomstmaten; (c) deze effecten behouden blijven tot minstens drie maanden na beëindiging van de behandeling; en (d) dergelijke effecten tevens geobserveerd werden buiten een onderzoeksetting. Er is sprake van '*sterke evidentie*' indien er aan minstens één van volgende criteria voldaan wordt: (a) matig tot hoogkwalitatieve evidentie dat de interventie tot een klinisch betekenisvol effect op symptomatologie of (b) functionele uitkomstmaten leidt. Indien daarentegen effecten gerapporteerd worden die niet klinisch betekenisvol zijn of afkomstig zijn van studies die niet aan hoge methodologische standaarden voldoen, spreekt men over '*zwakke evidentie*'. Hiertoe bieden we hieronder een overzicht van verschillende indicatoren voor de kwaliteit van een wetenschappelijke publicatie.

### **Evaluatie op basis van het wetenschappelijke tijdschrift**

Een eerste ruwe inschatting van de kwaliteit van de publicatie kan gemaakt worden op niveau van het wetenschappelijke tijdschrift waarin het artikel gepubliceerd werd. Zo kan nagegaan worden of het tijdschrift gebruik maakt van '*peerreview*', beoordeling van manuscripten door onafhankelijke experts. Daarnaast kan het aanzien van het tijdschrift geëvalueerd worden door de '*impactfactor*' op te zoeken en de positionering in de lijst van tijdschriften relevant voor het onderzoeksveld. De impactfactor vormt een indicator voor de frequentie waarmee wetenschappelijke werken uit het geselecteerde tijdschrift geciteerd worden. Deze waarde is echter erg relatief, tijdschriften uit grotere wetenschappelijke disciplines zoals oncologisch medisch onderzoek zullen vaker geciteerd worden en beschikken hierdoor mogelijk over een hogere impactfactor dan toptijdschriften in het veld van de klinische psychologie of psychiatrie. Beoordeling van de kwaliteit van wetenschappelijke tijdschriften op basis van de impactfactor van een tijdschrift dient dus steeds plaats te vinden binnen het eigen onderzoeksveld, waarbij men gebruik maakt van een *rangschikking* van impactfactoren. Deze informatie is mits abonnement beschikbaar via de Journal Citation Reports op WoS. Echter, er is heel wat kritiek op het gebruik van de impactfactor als maat voor kwaliteit.

### **Evaluatie op basis van de impact van het artikel binnen het onderzoeksveld**

Vervolgens kan de kwaliteit van een artikel ingeschat worden op basis van de impact van het artikel op de onderzoekslijn. Hiertoe kan gekeken worden naar het aantal *citaties* dat een bepaald artikel heeft. Deze informatie is terug te vinden via verschillende kanalen, waaronder WoS, GoogleScholar en ResearchGate. Deze laatste twee zijn interessant voor klinici, omdat ze citatierapporten van artikels vrij online beschikbaar stellen. Daarbij geldt dat – rekening houdend met de *recentheid* van een publicatie – een hoger aantal citaties door collega's in het werkveld gezien kan worden als indicator voor impact en erkenning van het artikel binnen het veld. Citaties kunnen echter tevens wijzen op kritiek op de studie. Een indicatie hiervan kan worden verkregen via ResearchGate, waarop men gratis full-text artikels kan raadplegen of verzoeken, inclusief de context waarin een artikel geciteerd werd.

### **Evaluatie op basis van het corpus van het manuscript**

Het corpus van het manuscript laat de meest verfijnde evaluatie van een artikel toe. Daarbij kunnen specifieke aanbevelingen gemaakt worden voor de verschillende luiken van de publicatie, achtereenvolgens te onderscheiden: de inleiding, methode-, resultaten- en discussiesectie. Dit dient zich tevens te weerspiegelen in het abstract van het artikel. Daarenboven is het belangrijk na te gaan of de auteurs van de publicatie strijdige belangen rapporteren ('*conflict of interest*'). Tabel 2 biedt een overzicht van aandachtspunten bij evaluatie van de kwaliteit van ongefilterde bronnen, zoals RCT's, aangezien deze een centrale positie innemen in behandelonderzoek.

#### **Inleiding.**

De inleiding van het artikel dient heldere *hypotheses* te presenteren die voldoende specifiek en logisch onderbouwd zijn. Deze zijn meestal terug te vinden op het einde van de inleiding. Daarbovenop zijn hypothesen van interventiestudies idealiter vooraf geregistreerd. *Pre-registratie* van een studie houdt in dat de onderzoeker, vooraleer over te gaan tot dataverzameling, de hypothesen, het onderzoeksopzet, de uitkomstmaten en analysestrategie vastlegt en deelt met de wetenschappelijke gemeenschap (bijv. via ClinicalTrials.Gov of Open Science Framework). Indien er sprake is van pre-registratie, wordt dit standaard vermeld

onderaan het abstract van een publicatie, aan het einde van de inleiding of het begin van de methodesectie, al dan niet in de vorm van een referentiecode.

### **Methode.**

Voor de methodesectie van een artikel zijn er tevens verschillende aandachtspunten die het evalueren van de kwaliteit van een interventiestudie – en de relevantie voor je onderzoeksvraag – kunnen faciliteren. Zo dient deze sectie een duidelijke beschrijving te bevatten van de *steekproef*, de *interventie(s)*, de *vergelijkingscondities*, *uitkomstmaten*, en het *verloop* van de studie. Deze dienen in die mate gedetailleerd te zijn dat een onafhankelijke onderzoeker het proces zou kunnen repliceren op basis van de beschrijvingen in het artikel. De *steekproefgrootte en karakteristieken* dienen afgestemd te zijn op de onderzoeksvraag. Zo wordt voor klinische vraagstellingen idealiter gebruik gemaakt van een passende klinische populatie. Daarnaast dient nagegaan te worden of de steekproefkarakteristieken overeenkomen met deze van uw individuele patiënt.

In het geval van behandelonderzoek is er best sprake van verschillende *vergelijkingscondities* die toelaten om te controleren voor ‘placebo’-effecten alsook effecten van tijd (bijv. spontaan herstel). Daarnaast dient beschreven te worden of participanten gerandomiseerd werden over deze vergelijkingcondities en hoe ‘*randomisatie*’ en eventuele ‘*blinding*’ van participanten plaatsvond. Bij randomisatie worden participanten random toegewezen aan de verschillende interventiecondities om op deze wijze eventuele variatie tussen de vergelijkingcondities voorafgaand aan de experimentele interventie uit te filteren. Blinding beoogt te voorkomen dat men op de hoogte is van de experimentele conditie en kan plaatsvinden op niveau van de participanten, behandelaars en onderzoekers. Idealiter is er sprake van een ‘double blind’ procedure, waarbij zowel de participanten als de overige betrokkenen bij het onderzoek niet op de hoogte zijn van de interventieconditie waartoe een participant gerandomiseerd werd. Dit laat toe om de invloed van (onbewuste) vertekeningen op de resultaten langs de zijde van de participanten én onderzoekers te minimaliseren. Afhankelijk van het type behandelonderzoek en de gebruikte vergelijkingcondities, is een double blind procedure echter niet steeds mogelijk.



Indien het gaat om klinische interventies dient tevens vermeld te worden hoe de *behandelaars getraind* werden en welke maatregelen genomen werden om te garanderen dat de behandelaars zich aan het *behandelprotocol* hielden. Andere relevante vragen zijn hoe omgegaan wordt met *uitval* doorheen de studie (bijv. intention-to-treat analyses; werd er rekening gehouden met participanten die niet meer opdaagden of de interventie niet voltooiden?) en wat de *psychometrische kwaliteiten* zijn van de uitkomstmaten. In het geval dat de studie gebruik maakt van pre-registratie, kan ook nagegaan worden of de condities en uitkomstmaten die gerapporteerd worden in de publicatie overeenstemmen met deze in de pre-registratie.

### **Resultatensectie en discussie.**

Basiskennis rond statistische proeven en principes is van belang bij zowel het beoordelen van de kwaliteit van de resultatensectie als de daarbij horende interpretaties in de discussiesectie. Hiertoe voorzien de opleidingen klinische psychologie, orthopedagogie, en geneeskunde een aantal statistiekvakken wat de hulpverlener kan bijstaan in het herkennen van veelvoorkomende statistische vraagstellingen, het identificeren van de bijhorende statistische proeven en het interpreteren van de resultaten. Om daarbij de statistische besluitvorming te kunnen volgen, is het belangrijk dat *descriptieve waarden* zoals gemiddelden en standaarddeviaties voor de verschillende condities weergegeven worden. Daarnaast dienen ook de *grootte van effecten* (effect size) en de *stabiliteit* van de bevindingen gerapporteerd te worden (bijv. betrouwbaarheidsintervallen). Een veelgebruikte maat voor effectgrootte is Cohen's *d* (Cohen, 1988), waarbij een score van 0.20 – 0.49 overeenstemt met een klein effect, 0.50 – 0.79 met een medium effect en er vanaf 0.80 gesproken wordt van een groot effect.

Andere belangrijke vragen in deze context zijn of er gecontroleerd werd voor *meervoudig toetsen* en of de studie voldoende *power* had om de effecten in kaart te brengen. Zo wordt de kans op het vinden van statistisch significante bevindingen groter naarmate meer toetsen uitgevoerd worden. Hiervoor kan gecontroleerd worden door gebruik te maken van aangepaste en conservatievere normen voor statistische significantie. Wat power betreft, of de kans om een bestaand effect te observeren als zijnde statistisch significant, dient men zich af te vragen of gebruik gemaakt werd van een voldoende grote steekproef om het verwachte effect

te observeren. Idealiter wordt hiertoe een poweranalyse in de methodesectie of eventuele pre-registratie gerapporteerd. Omgekeerd is het goed om te weten dat de kans op statistisch significante bevindingen tevens groter wordt naarmate er gebruik gemaakt wordt van zeer grote steekproeven. Deze zijn echter niet noodzakelijk klinisch betekenisvol. Voor een klinische vraagstelling wordt idealiter tevens gekeken naar de *klinische waarde* van een bevinding. In de psychiatrische en klinisch psychologische literatuur worden hiertoe verschillende maten naar voor geschoven, waaronder '*number needed to treat*'. Deze waarde geeft het aantal patiënten weer die met een interventie behandeld dienen te worden vooraleer dit een positief effect oplevert, zoals het uitblijven van een ongunstige uitkomst. In geval van pre-registratie kan opnieuw nagegaan worden of de gepresenteerde analyses in het manuscript overeenstemmen met de pre-registratie.

Bij kritische lezing van de discussiesectie is het vervolgens van belang om na te gaan of de *interpretaties* en conclusies gemaakt kunnen worden op basis van het gebruikte onderzoeksdesign (bijv. uitspraken rond *causaliteit* zijn enkel mogelijk in geval van experimentele designs) en of deze overeenstemmen met de gerapporteerde resultaten. Hierbij dient opnieuw de vraag gesteld te worden naar de toepasbaarheid voor de individuele patiënt.

### **Bijkomende hulpbronnen bij kritische beoordeling van publicaties**

Kritische evaluatie van wetenschappelijke artikels is een vaardigheid die initieel heel wat oefening en investering vergt van de clinicus. Echter, mits voldoende oefening laat dit snelle en efficiënte evaluatie van wetenschappelijke bevindingen toe. Een aantal hulpmiddelen kunnen het verder ontwikkelen van deze vaardigheid faciliteren (zie Tabel 3).

### **Conclusie**

De vraag naar evidence-based practice confronteert hulpverleners met de nood aan specifieke kennis en vaardigheden gericht op het kritisch beoordelen van wetenschappelijke evidentie. Dit artikel beoogt concrete richtlijnen te bieden die toelaten een eerste inschatting te maken van de kracht alsook de kwaliteit van de evidentie op niveau van het type bron, het tijdschrift en het specifieke artikel. Daarbij focussen we ons op evaluatie van ongefilterde

bronnen zoals RCT's aangezien deze een centrale positie innemen in behandelonderzoek. Daarnaast stellen we verschillende hulpmiddelen voor die hulpverleners kunnen gebruiken bij evaluatie van andere bronnen.

## Referenties

- APA Presidential Task Force on Evidence-Based Practice (2006). Evidence-Based Practice in Psychology. *American Psychologist*, 61, 271-285. doi: 10.1037/0003-066X.61.4.271
- Australian National Health and Medical Research Council (2015). NHMRC levels of evidence. Retrieved from <https://lifeinthefastlane.com/cccl/levels-and-grades-of-evidence/>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.
- Guyatt, G. H., Oxman, A. D., Vist, G. E., Kunz, R., Falck-Utter, Y., Alonso-Coello, P., & Schünemann, H. J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336, 924-926. doi: 10.1136/bmj.39489.470347
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gotzsche, P. C., Devereaux, P. J., ... Altman, D. G. (2010). CONSORT 2010 Explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *British Medical Journal*, 340:28. doi: 10.1136/bmj.c869
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., . . . Stewart, L. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4,, 1-9.
- Oxford Centre for Evidence-Based Medicine Levels of Evidence Working Group (2011). The Oxford 2011 levels of evidence. Retrieved from [www.cebm.net/wp-content/uploads/2014/06/CEBM-levels-of-evidence-2.1.pdf](http://www.cebm.net/wp-content/uploads/2014/06/CEBM-levels-of-evidence-2.1.pdf)
- Tolin, D. F., McKay, D., Forman, E. M., Klonsky, E. D., & Thombs, B. D. (2015). Empirically supported treatment: Recommendations for a new model. *Clinical Psychology Science and Practice*, 22, 317-338. doi: 10.1111/cpsp.12122

Tabel 1. Kracht van evidentie op basis van type bron

<b>Niveau</b>	<b>Type bron</b>
Niveau 1	Wetenschappelijke evidentie verkregen uit een systematische review bestaande uit alle beschikbare relevante RCT's (o.a. klinische richtlijnen, meta-analyses, review artikels)
Niveau 2	Wetenschappelijke evidentie verkregen uit minstens één RCT
Niveau 3a	Wetenschappelijke evidentie verkregen uit klinische trials waarin gebruik gemaakt wordt van pseudo-randomisatie (d.i., afwijkend van de random allocatie per conditie)
Niveau 3b	Wetenschappelijke evidentie bekomen uit cohort studies (bijv. voor onderzoek rond etiologische factoren), case-control studies, of quasi-experimenten zoals onderbroken time-series designs met een controleconditie
Niveau 3c	Wetenschappelijke evidentie bekomen uit studies met een historische controlegroep, twee of meer studies zonder controleconditie, of onderbroken time-series designs met een parallelle controleconditie
Niveau 4	Wetenschappelijke evidentie bekomen uit gevalsstudies

Tabel 2. Evaluatie van wetenschappelijke artikels

Niveau	Aandachtspunten
<i>Abstract</i>	<ol style="list-style-type: none"> <li>1. Beknopte rationale, overzicht voornaamste methodologische kenmerken, hoofdbevindingen en passende interpretatie/conclusie</li> <li>2. Eventuele vermelding van pre-registratie (identificatiecode)</li> </ol>
<i>Inleiding</i>	<ol style="list-style-type: none"> <li>3. Duidelijke rationale</li> <li>4. Gaat het om exploratief vs. confirmatorisch onderzoek?</li> <li>5. Heldere onderzoeksvragen en bijhorende hypothesen</li> <li>6. Conform gegevens van eventuele pre-registratie</li> </ol>
<i>Methode</i>	<ol style="list-style-type: none"> <li>7. Laat replicatie van de studie toe</li> <li>8. Bevat uitgebreide beschrijving van steekproefgrootte, steekproefkarakteristieken, rekruteringsstrategie en uitval doorheen de studie</li> <li>9. Bevat uitgebreide beschrijving van diagnostische instrumenten/uitkomstmaten en psychometrische kwaliteiten</li> <li>10. Evalueer de relevantie van de gebruikte steekproef voor jouw individuele patiënt</li> <li>11. Bevat uitgebreide beschrijving van vergelijkingscondities</li> <li>12. Randomisatie: Worden participanten random toegewezen aan de interventiecondities?</li> <li>13. Controlegroep: Laten de vergelijkingscondities toe te controleren voor placebo-effecten of effecten van tijd?</li> <li>14. Blinding: Is er sprake van 'blinding' en bij welke partijen?</li> <li>15. Voor psychotherapeutische interventies: Bevat een beschrijving van hoe behandelaars getraind werden alsook wat ondernomen wordt om te evalueren of behandelaars zich aan het protocol hielden</li> <li>16. Conform gegevens van eventuele pre-registratie of duidelijke argumentatie</li> </ol>

voor afzien van de origineel geplande methode / procedure

- Resultaten*
- 17. Logische opbouw en selectie van toetsen
  - 18. Wordt rekening gehouden met eventuele schending van assumpties van statistische toetsen?
  - 19. Worden descriptieve waarden voor de verschillende condities weergegeven (gemiddelden, standaarddeviaties)?
  - 20. Worden grootte van effecten en betrouwbaarheidsintervallen gerapporteerd?
  - 21. Wordt uitval in rekening gebracht (bijv. via intention-to-treat analyses)?
  - 22. Bevat de studie voldoende power voor de uitgevoerde analyses en verwachte effecten (cfr. steekproefgrootte)?
  - 23. Werd eventueel gecontroleerd voor meervoudig toetsen?
  - 24. Maak onderscheid tussen statistische versus klinische significantie
  - 25. Conform gegevens van eventuele pre-registratie of duidelijke argumentatie voor afzien van de origineel voorziene analysestrategie

- Discussie*
- 26. Zijn interpretaties in lijn met de bevindingen uit de resultatensectie?
  - 27. Kunnen de interpretaties gemaakt worden o.b.v. het onderzoeksdesign?
  - 28. Bevindingen gekaderd binnen de bredere literatuur
  - 29. Kritische basishouding met oog voor beperkingen van de studie en aanbevelingen voor toekomstig onderzoek
  - 30. Conclusie afgestemd op de resultaten

- Verklaring*
- 31. Hou rekening met eventuele strijdige belangen
- strijdige*  
*belangen*
-

Tabel 3. Hulpmiddelen

Type	Inhoud
Trainingsmodules	<p>Het Centrum voor Evidence-Based Behavioral Practice biedt handleidingen en oefeningen voor evidence-based werken in de praktijk (<a href="http://ebbp.org/training.html">http://ebbp.org/training.html</a>)</p> <p>National Registry of Evidence-based Programs and Practices; hulpmiddelen om te leren selecteren tussen en stapsgewijs implementeren van evidence-based behandelingen (<a href="https://nrepp-learning.samhsa.gov/samhsa-evidence-based-practice-kits">https://nrepp-learning.samhsa.gov/samhsa-evidence-based-practice-kits</a>)</p>
Checklists	<p>Kritisch beoordelen van informatie via het Critical Appraisal Skills Programme (CASP; <a href="http://www.casp-uk.net/checklists">http://www.casp-uk.net/checklists</a>)</p> <p>Consolidated Standards of Reporting Trials (CONSORT criteria; Moher et al., 2010) ter beoordeling van de kwaliteit van RCT's</p> <p>Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA richtlijnen; Moher et al., 2015) ter beoordeling van reviews/meta-analyses</p> <p>Appraisal of Guidelines for Research and Evaluation instrument (<a href="http://www.agree.trust.org">www.agree.trust.org</a>) ter evaluatie van behandelrichtlijnen</p>